1. Elementary Statistics: Introduction
2. Lecture Videos
    1. Video Lecture 1: Sampling and Data
    2. Video Lecture 2: Descriptive Statistics
    3. Video Lecture 3: Linear Regression and Correlation
    4. Video Lecture 4: Probability Topics
    5. Video Lecture 5: Discrete Distributions
    6. Video Lecture 7: The Normal Distribution
    7. Video Lecture 8: The Central Limit Theorem
    8. Video Lecture 9: Confidence Intervals
    9. Video Lecture 10: Hypothesis Testing with a Single Mean
    10. Video Lecture 11: Hypothesis Testing with Two Means
    11. Video Lecture 12: The Chi-Square Distribution
3. Practice Exams and Quizzes
    1. Skills Practice Exams
        1. Elementary Statistics: Skills Practice Exam 1- Lessons 1, 2, 12
        2. Skills Practice Exam 2: Chapters 3, 4, 5, & 6
        3. Elementary Statistics: Skills Practice Exam 3- Lessons 7,8, 9, 10
    2. Practice Final Exams
        1. Elementary Statistics: Exam 1: Lessons 1 & 2
        2. Elementary Statistics: Exam 2: Lessons 3 & 4
        3. Elementary Statistics: Exam 3: Lessons 5, 6 & 7
        4. Elementary Statistics: Exam 4: Lessons 8, 9 & 10
        5. Elementary Statistics: Exam 5: Lesson 11
        6. Exam 6: Chapter 12
4. Calculator Instructions
    1. Elementary Statistics: Calculator Instructions and Resources

Elementary Statistics: Introduction

The following modules are based on the award-winning *Elementary Statistics* online course by authors Barbara Illowsky and Susan Dean. The content presented here was designed to be used as a complementary resource with their Collaborative Statistics textbook/collection.

Note: The source documents for this collection can be found at http://sofia.fhda.edu/gallery/statistics/index.html.

Course Management

The Course Syllabus provides instructors a basic framework for teaching this material to their students. Thi document is intended to serve as a starting point; instructors should use this document as a foundation for creating a learning experience customized to meet their students unique needs.

Video Lectures

As a part of their award-winning online course, the authors have provided a number of video lectures. These half-hour segments can be used for self-study or as a complement to the Collaborative Statistics textbook. The authors also provide videos instructing students on the use of the TI-83 calculator as used in the textbook and course activities and exercises.

Lecture Videos

- Chapter 1: Sampling and Data
- Chapter 2: Descriptive Statistics
- Chapter 3: Probability Topics
- Chapter 4: Discrete Distributions
- Chapter 5: Continuous Random Variables
- Chapter 6: The Normal Distribution

- Chapter 7: The Central Limit Theorem
- Chapter 8: Confidence Intervals
- Chapter 9: Hypothesis Testing - Single Mean and Single Proportion
- Chapter 10: Hypothesis Testing - Two Means, Two Proportions, Paired Data
- Chapter 11: The Chi-Square Distribution
- Chapter 12: Linear Regression and Correlation

TI-83 Calculator Video Tutorials

- TI-83 Calculator Tutorial, Part 1
- TI-83 Calculator Tutorial, Part 2

Practice Exams, Problem Sets, and Quizzes

A number of practice tests and problem sets are provided for student self-evaluation and to provide opportunities for students to practice key concepts introduced throughout the course. Solutions to these exercises are provided as feedback to aid student retention and understanding. These problem sets may be used as homework assignments or self-directed study aids.

Skills Practice Exams

- Skills Practice Exam 1: Chapters 1, 2, & 12
- Skills Practice Exam 2: Chapters 3, 4, 5, & 6
- Skills Practice Exam 3: Chapters 7, 8, 9, & 10

Practice Final Exams

- Practice Final Exam 1: Chapters 1 & 2
- Practice Final Exam 2: Chapters 3 & 4
- Practice Final Exam 3: Chapters 5, 6, & 7
- Practice Final Exam 4: Chapters 8, 9, & 10
- Practice Final Exam 5: Chapter 11
- Practice Final Exam 6: Chapter 12

In addition to the problem sets provided above, the following multiple-choice quizzes are provided as resources for instructors. These modules can

be used as assignments or as templates for classroom assessments. Answers to these items are not provided.

Quizzes

- Chapter 1: Sampling and Data
- Chapter 2: Descriptive Statistics
- Chapter 3: Probability Topics
- Chapter 4: Discrete Distributions
- Chapter 5: Continuous Random Variables
- Chapter 6: The Normal Distribution
- Chapter 7: The Central Limit Theorem
- Chapter 8: Confidence Intervals
- Chapter 9: Hypothesis Testing - Single Mean and Single Proportion
- Chapter 10: Hypothesis Testing - Two Means, Two Proportions, Paired Data
- Chapter 11: The Chi-Square Distribution
- Chapter 12: Linear Regression and Correlation

Calculator Instructions

The following module contains a number of resources related to the TI-83 calculator and ways it can be used with the Collaborative Statistics textbook and curriculum. This resource addresses many different function on the calculator, including calculation of the outliers, discrete mean, standard deviation, and random numbers.

- View the TI-83 Calculator Resources

Video Lecture 1: Sampling and Data

This video lecture is designed as a supplement to the Sampling and Data chapter of Collaborative Statistics (col10522) by Barbara Illowsky and Susan Dean. In this chapter, students can expect to learn how to recognize and differentiate between key terms in statistics, apply various types of sampling methods to data collections, and create and interpret frequency tables.

https://www.youtube.com/embed/viXNnZO8X9k

Video Lecture 2: Descriptive Statistics

This video lecture is designed as a supplement to the Descriptive Statistics chapter of Collaborative Statistics (col10522) by Barbara Illowsky and Susan Dean. In this chapter, students can expect to learn how to display data graphically and interpret graphs; recognize, describe and calculate the measures of location of data such as quartiles and percentiles; recognize, describe, and calculate the measures of the center of data; and recognize, describe, and calculate the measures of the spread of data.

https://www.youtube.com/embed/X_VJfdMTpLI

Video Lecture 3: Linear Regression and Correlation

This video lecture is designed as a supplement to the Linear Regression and Correlation chapter of Collaborative Statistics (col10522) by Barbara Illowsky and Susan Dean. In this chapter, students can expect to learn how to discuss basic ideas of linear regression and correlation; create and interpret a line of best fit; and calculate and interpret the correlation coefficient and outliers.

https://www.youtube.com/embed/hbKcpJaPknI

Video Lecture 4: Probability Topics

This video lecture is designed as a supplement to the Probability Topics chapter of Collaborative Statistics (col10522) by Barbara Illowsky and Susan Dean. In this chapter, students can expect to learn how to understand and use the terminology of probability; determine whether two events are mutually exclusive or independent; calculate probabilities using the Addition Rules and Multiplication Rules; and construct and interpret Contingency Tables, Venn Diagrams, and Tree Diagrams.

https://www.youtube.com/embed/fpLlgdWfcE0

Video Lecture 5: Discrete Distributions

This video lecture is designed as a supplement to the Discrete Distributions of Collaborative Statistics (col10522) by Barbara Illowsky and Susan Dean. In this chapter, students can expect to learn how to recognize and understand discrete probability distribution functions; calculate and interpret expected values; recognize and apply the binomial probability distribution, the Poisson probability distribution, the geometric probability distribution, and the hypergeometric probability distribution; and classify discrete word problems by their distributions.

https://www.youtube.com/embed/vy_eY1H_D0U

Video Lecture 7: The Normal Distribution

This video lecture is designed as a supplement to the Normal Distribution chapter of Collaborative Statistics (col10522) by Barbara Illowsky and Susan Dean. In this chapter, students can expect to learn how to recognize the normal probability distribution and the standard normal probability distribution and apply each appropriately, and compare normal probabilities by converting to the standard normal distribution.

https://www.youtube.com/embed/OqorvQXzemY

Video Lecture 8: The Central Limit Theorem

This video lecture is designed as a supplement to the Central Limit Theorem chapter of Collaborative Statistics (col10522) by Barbara Illowsky and Susan Dean. In this chapter, students can expect to learn how to recognize the Central Limit Theorem problems; classify continuous word problems by their distributions, and apply and interpret the Central Limit Theorem for Averages and the Central Limit Theorem for Sums.

https://www.youtube.com/embed/qZFDvprWLN0

Video Lecture 9: Confidence Intervals

This video lecture is designed as a supplement to the Confidence Intervals chapter of Collaborative Statistics (col10522) by Barbara Illowsky and Susan Dean. In this chapter, students can expect to learn how to calculate and interpret confidence intervals for one populations average and one population proportion; interpret the student-t probability distribution as the sample size changes; and discriminate between problems applying the normal and then student-t distributions.

https://www.youtube.com/embed/Gybpz9GdxAU

Video Lecture 10: Hypothesis Testing with a Single Mean
This video lecture is designed as a supplement to the Hypothesis Testing: Single Mean and Single Proportion chapter of Collaborative Statistics (col10522) by Barbara Illowsky and Susan Dean. In this chapter, students can expect to learn how to differentiate between Type I and Type II Errors; describe hypothesis testing in general and in practice; and conduct and interpret hypothesis tests for single population means, known and unknown population standard deviation, and also proportions.

Note: A captioned version of this video is currently unavailable. This will be updated shortly.
https://www.youtube.com/embed/gpL0mAZ_JfE

Video Lecture 11: Hypothesis Testing with Two Means
This video lecture is designed as a supplement to the Hypothesis Testing: Two Means, Two Populations, Paired Data chapter of Collaborative Statistics (col10522) by Barbara Illowsky and Susan Dean. In this chapter, students can expect to learn how to classify hypothesis tests by type; conduct and interpret hypothesis tests for two population means with standard deviations that are both known and unknown; conduct and interpret hypothesis tests for two population proportions; and conduct and interpret hypothesis tests for matched or paired samples.
https://www.youtube.com/embed/KSgDoWWZevU

Video Lecture 12: The Chi-Square Distribution

This video lecture is designed as a supplement to the Chi-Square Distribution chapter of Collaborative Statistics (col10522) by Barbara Illowsky and Susan Dean. In this chapter, students can expect to learn how to interpret the chi-square probability distribution as the sample size changes; conduct and interpret chi-square goodness-of-fit hypothesis tests; conduct and intepret chi-square test of independence tests; and conduct and interpret chi-square single variance hypothesis tests.

https://www.youtube.com/embed/S5jbIw0SI2s

Elementary Statistics: Skills Practice Exam 1- Lessons 1, 2, 12

Questions 1 – 4 use the following table and shows the lengths (in minutes) of 64 international phone calls using a prepaid calling card.

Length of call (minutes)	**Frequency**	**Relative frequency**	**Cum. Relative Freq.**
4	25	0.3906	
14	15		
24	10	0.1563	
34	9	0.1406	
44	4	0.0625	
54	1	0.0156	1.00

Exercise:

Problem: The histogram of this data looks to be:

- **A**Skewed right
- **B**Skewed left
- **C**Symmetrical

Solution:

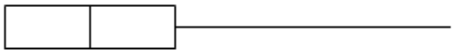
A

Exercise:

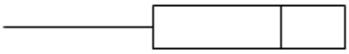
Problem: Which of the following box plots most accurately displays the data?

- **A**

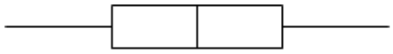
- **B**



- **C**



- **D**



Solution:

B

Exercise:

Problem: What percent of telephone calls were more than 24 minutes?

- **A**15.63%
- **B**21.88%
- **C**62.5%
- **D**78.13%

Solution:

B

Exercise:

Problem: Find the 80th percentile.

- **A**14

- **B**24
- **C**34
- **D**70

Solution:

C

Exercise:

Problem:

What can be said about a set of data when its standard deviation is small (but not zero)?

- **A**The data are far apart.
- **B**All of the data have the same value.
- **C**The mean of the data can never be zero.
- **D**The data are close together.

Solution:

D

Questions 6 and 7, refer to the following: A sample of students was taken to determine pulse rate. The data is shown:

Pulse rate-beats per minute	54	58	65	68	72	76	80	90	98
Frequency (# of Students)	1	3	6	8	5	3	8	4	2

Exercise:

Problem: Find the median and mode(s).

- **A**72, 68 and 80
- **B**72 and 80
- **C**76 and 68
- **D**76, 68 and 80

Solution:

A

Exercise:

Problem:

Out of the entire college population of 24,000 students, approximately what percent of students are expected to have a pulse rate of 65?

- **A**6
- **B**40
- **C**9
- **D**15

Solution:

D

For questions 8 and 9, refer to the following: By determining the average number of people in a car using the "Diamond" Carpool Lane, the California Highway Patrol is trying to decide if the number of people in a car using the "Diamond" Carpool Lane should be increased from 2 to 3.

Exercise:

Problem:

The average number of people per car for all cars using the "Diamond" Carpool Lane is called the:

- **A**parameter
- **B**data
- **C**variable

- **D**statistic

Solution:

D

Exercise:

Problem: The number of people in 1 car is called the:

- **A**parameter
- **B**data
- **C**variable
- **D**statistic

Solution:

C

Exercise:

Problem:

I toss a fair coin a large number of times. Assuming the tosses are independent, which of the following is true?

- **A**Once the number of flips is large enough, the number of heads will always be exactly half of the total number of tosses. For example, after 10,000 tosses I should have exactly 5,000 heads.
- **B**The proportion of heads will be about 1/2 and this proportion will tend to get closer to 1/2 as the number of tosses increases.
- **C**As the number of tosses increases, any long run of heads will be balanced by a corresponding run of tails so that the overall proportion of heads is exactly 1/2.
- **D**All of the above.

Solution:

B

Questions 11 - 13, refer to the following: A sample of twenty people went on a cruise to Alaska. Their two-week weight gain is shown below (a weight loss is shown by a negative number.)

Weight Gain	Frequency
-2	4
0	5
2	8
5	2
9	1

Exercise:

Problem: The middle 50% of the data is between _____ and _____.

- **A** 0 and 2
- **B** 0 and 9
- **C** 2 and 9
- **D** 2 and 2

Solution:

A

Exercise:

Problem: Find the average weight gain (in pounds).

- **A** 1.35
- **B** 2.74
- **C** 2

- **D**There is not enough information.

Solution:

A

Exercise:

Problem:

What weight gain is 3 standard deviations above the mean (in pounds)?

- **A**4.05
- **B**8.19
- **C**9.57
- **D**There is not enough information

Solution:

A

For questions 14 - 19, use the following information: Kim, a personal trainer, was interested in whether or not there was a linear relationship between the number of visits her clients made to the gym each week and the average amount of time her clients exercised per visit. She took the following data.

Client	1	2	3	4	5	6
Number of visits per week	1	3	4	2	3	5
Average time spent exercising per visit (hours)	2	1.5	1	2	2	0.30

Exercise:

Problem: The line that best fits the data is:

- **A** $\hat{y} = -0.44 + 2.62x$
- **B** $\hat{y} = 0.44 + 2.62x$
- **C** $\hat{y} = 2.62 + 0.44x$
- **D** $\hat{y} = 2.62 - 0.44x$

Solution:

A

Exercise:

Problem: Is the correlation coefficient significant?

- **A** Yes
- **B** No
- **C** It might be.
- **D** Not enough information is given.

Solution:

A

Exercise:

Problem:

Using the best fit line, estimate the average time spent exercising per visit for 4 visits per week.

- **A** 2 hours
- **B** 0.86 hours
- **C** 1 hour
- **D** 10.04 hours

Solution:

B

Exercise:

Problem:

Kim used the best fit line to estimate the average time spent exercising per visit for her client Toby who visited the gym 7 times per week. Does the least squares line give an accurate estimate?

- **A**Yes
- **B**No
- **C**Maybe
- **D**Not enough information is given.

Solution:

B

Exercise:

Problem: If the correlation coefficient is –1, which answer is correct?

- **A**The slope of the best fit line is positive.
- **B**The slope of the best fit line is –1.
- **C**The data fit exactly on a line with positive slope.
- **D**The data fit exactly on a line with negative slope.

Solution:

D

Exercise:

Problem: A scatter plot shows:

- **A**the direction and strength of a relationship between the independent and dependent variables.
- **B**that there is a linear relationship between the independent and dependent variables.
- **C**how you can predict the dependent variable knowing the independent variable.
- **D**nothing. The line of best fit is what is important.

Solution:

A

Exercise:

Problem:

Suppose we are interested in the average grade on the first math 10 test taken by all students at De Anza during the spring 2002 quarter. We randomly choose 3 students from each of the spring 2002 Math 10 classes as our sample. This sampling technique is

- **A**systematic
- **B**cluster
- **C**stratified
- **D**convenience

Solution:

B

Exercise:

Problem:

Tracy works at a local indoor soccer arena. He is interested in the proportion of people entering the area who spend money in the arena store. One night while he is working, Tracy counts the first 20 people who buy goods in the arena store.

- **A**systematic
- **B**cluster
- **C**stratified
- **D**convenience

Solution:

D

Skills Practice Exam 2: Chapters 3, 4, 5, & 6

Questions 1 – 2 refer to the following: Below is the probability distribution function for the number of high school years that students at a local high school play on a sports team.

X	$P(X = x)$
0	0.32
1	0.12
2	
3	0.18
4	0.14

Exercise:

Problem: What is the probability that $X = 2$?

- **A**0.24
- **B**0.76
- **C**0.32
- **D**Cannot determine

Solution:

A

Exercise:

Problem:

Over the long run, the average number of years that we would expect students at this high school to play on a sports team is:

- **A**0
- **B**1.7
- **C**2
- **D**2.6

Solution:

B

Exercise:

Problem:

According to the 2000 United States Census, 12.3% of the population is Black or African American. The probability that a randomly selected U. S. resident is NOT Black or African American is:

- **A**0.123
- **B**0.877
- **C**0.754
- **D**Cannot determine

Solution:

B

Exercise:

Problem:

Assume the statistics final is a multiple choice test with 40 questions. Each question has four choices with one correct answer per question. If you were to randomly guess on each of the questions, what is the probability of getting exactly the expected number of correct answers?

- **A**0.5839
- **B**0.5605
- **C**0.25
- **D**0.1444

Solution:

D

Exercise:

Problem: In an exponential distribution, the mean is larger than the median.

- **A**true
- **B**false

Solution:

A

Exercise:

Problem:

In Fall 1999, students in one Math 10 section determined that the length of movies at the cinema was normally distributed with a mean of 148 minutes and a standard deviation of 19 minutes. Find the third quartile and interpret it.

- **A**75 minutes; Three-fourths of the movie lengths fall below 75 minutes.
- **B**160.8 minutes; Three-fourths of the movie lengths fall below 160.8 minutes.
- **C**160.8; Three-fourths of the movies last 160.8 minutes.
- **D**75 minutes; Three-fourths of the movies last 75 minutes.

Solution:

B

Exercise:

Problem: Which of the following is FALSE about data that follows the normal distribution?

- **A**The mean is the same as the mode.
- **B**The standard deviation is the same as the mean.
- **C**The median is the same as the mode.
- **D**Most data is within 3 standard deviations of the median.

Solution:

B

Exercise:

Problem:

The graph showing the age of getting a driver's license in California starts and peaks at age 16, and decreases from there. This shape most closely resembles what type of distribution?

- **a** Normal
- **b** Binomial
- **c** Uniform
- **d** Exponential

Solution:

D

Use the following information for questions 9 and 10. The amount of time that a randomly chosen 6th grade student spends on homework per week is uniformly distributed from 30 to 120 minutes.

Exercise:

Problem:

What is the probability that a randomly chosen 6th grade student spends at least 60 minutes per week on homework knowing that he/she will spend at most 80 minutes per week on homework?

- **a** 1.20
- **b** 0.6667
- **c** 0.2222
- **d**0.4

Solution:

D

Exercise:

Problem:

What is the expected amount of time that a randomly chosen 6th grade student spends on homework per week?

- **a** 45 minutes
- **b** 60 minutes
- **c** 30 minutes
- **d**75 minutes

Solution:

D

Use the following information for questions 11 and 12. The length of time a randomly chosen 9-year old child spends playing video games per day is approximately exponentially distributed with a mean equal to 2 hours.
Exercise:

Problem: Find the probability that a randomly chosen 9-year old will play video games at most 3 hours.

- **a** 0.7769
- **b** 0.9975
- **c** 0.0025
- **d**0.2231

Solution:

A

Exercise:

Problem: 70% of 9-year old children will play video games per day for at most how long?

- **a** 0.60 hours
- **b** 2.41 hours
- **c** 0.71 hours
- **d**Cannot determine

Solution:

B

Use the following information for questions 13 and 14. Research has shown that studying improves a student's chances to 80% of selecting the correct answer to a multiple choice question. A multiple choice test has 15 questions. Each question has 4 choices.
Exercise:

Problem: What is the distribution for the number of questions answered correctly when a student studies?

- **a** B(15, 0.80)
- **b** B(15, 0.25)
- **c** P(15)
- **d** P(6)

Solution:

A

Exercise:

Problem:

Suppose that a student does not study for the test but randomly guesses the answers. What is the probability that the student will answer 7 or 8 questions correctly?

- **a** 0.2951
- **b** 0.0524
- **c** 0.0131
- **d**Cannot determine

Solution:

B

Exercise:

Problem:

A downtown hotel determined the probability of finding X taxicabs waiting outside the hotel anytime between 5 PM and midnight. The information is shown in the table.

X	$P(X)$
1	0.0667
2	0.1331
3	0.2000
4	0.2667
5	0.3333

What is the average number of taxicabs that are expected to be waiting outside the hotel anytime between 5 PM and midnight?

- **a** 3.7
- **b** 3
- **c** 0
- **d** 15

Solution:

A

Exercise:

Problem:

. During the registration period for a new quarter, the De Anza College Registrar's Office processes approximately 75 applications per hour, on the average. What is the probability that it will process more than 80 applications for a randomly chosen hour? (This is a Poisson problem. If you did not cover the Poisson Distribution, then skip this problem.)

- **a** 0.0379
- **b** 0.2589
- **c** 0.7411
- **d**0.0248

Solution:

B

Questions 17 - 19 refer to the following: P(T) = 0.69 P(S) = 0.5, P(S|T) = 0.5
Exercise:

Problem: Events S and T are:

- **a** mutually exclusive
- **b** independent
- **c** mutually exclusive and independent
- **d** neither mutually exclusive nor independent

Solution:

B

Exercise:

Problem: Find P(S AND T)

- **a** 0.3450
- **b** 0.2500
- **c** 0.6900
- **d** 1

Solution:

A

Exercise:

Problem: Find P(S OR T)

- **a** 0.6900
- **b** 1.19
- **c** 0.8450
- **d** 0

Solution:

C

Exercise:

Problem:

Based on data from the US Census Bureau the average age of US residents is 36.31 with a standard deviation of 21.99. The data is normally distributed. The notation for the distribution is:

- **a** X ~ N(36.31, 21.99)
- **b** X ~ N(21.99, 36.31)
- **c** X ~ B(36.31, 22)
- **d** X ~ U(0, 36.31)

Solution:

A

Exercise:

Problem: In a binomial distribution we:

- **a** count the number of successes until a failure is obtained
- **b** count the number of trials until a success is obtained
- **c** count the number of successes in a finite number of trials
- **d** count the number of trials until the number of successes equals the number of failures

---

Solution:

C

Exercise:

Problem:

Certain stocks have a probability of 0.6 of returning a $100 profit. They also have a probability of 0.4 of having a loss of $300. Over the long run, what is the best thing to do to maximize your profit, and why?

- **a** Invest in the stocks because there is a greater probability of making money than losing money.
- **b** Do not invest in the stocks because the dollar amount for each loss is greater than the dollar amount for each gain.
- **c** Invest in the stocks because making $100 per stock is preferred to losing $300 per stock.
- **d** Do not invest in the stocks because the expected value is a loss.

Solution:

D

Questions 23 - 27 refer to the following table (data from the Institutional Research department of the Foothill-De Anza Community College District for De Anza College).

			Black	Hispanic	White	Undeclared	Total

	American Indian	Asian/Pacific Islander					
Administrator	0	3	5	5	21	0	34
Staff	1	35	21	30	201	16	304
Faculty	3	58	14	45	141	17	278
Total	4	96	40	80	363	33	616

Suppose that one De Anza College employee is randomly selected.

Exercise:

Problem: Find P (the employee is an Administrator)

- **a** 278/34
- **b** 304/616
- **c**34/616
- **d**80/616

Solution:

C

Exercise:

Problem: Find P (the employee is Faculty AND American Indian)

- **a** 382/616
- **b** 3/616
- **c** 3/4
- **d**3/278

Solution:

B

Exercise:

Problem: Find P (employee is Staff OR Hispanic)

- **a** 384/616
- **b** 80/616
- **c** 304/616
- **d**354/616

Solution:

D

Exercise:

Problem: Find P (employee is an Administrator GIVEN the employee is Black)

- **a** 40/616
- **b** 5/34
- **c** 5/616
- **d** 5/40

Solution:

D

Exercise:

Problem: Being an Administrator and an American Indian are

- **a** mutually exclusive events
- **b** independent events
- **c** mutually exclusive and independent events
- **d** neither mutually exclusive nor independent events

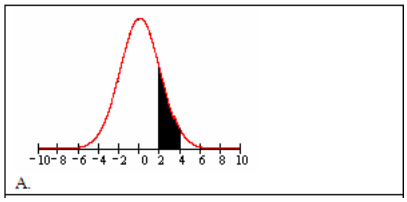
Solution:

A

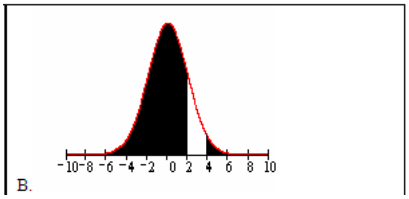
Exercise:

Problem: P(X$gt; 4 or X < 2) is best described by which of the following graphs?

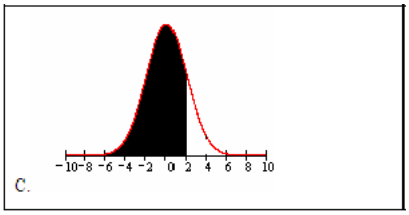
- **a**



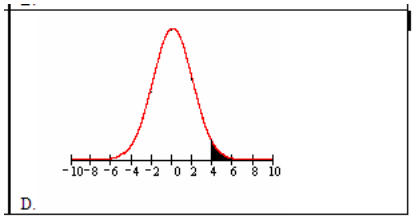
- **b**



- **c**



- **d**



Solution:

B

Questions 29 - 32 refer to the following: When a customer calls the "Help Line" at ABC Computer Software Co., the amount of time that a customer must wait "on hold" until somebody answers the line and helps the customer follows an exponential distribution with mean of 7.5 minutes.

Exercise:

Problem: What is the probability that a customer waits more than 10 minutes to receive help?

- **a** 0.2636
- **b** 0.75
- **c** 0.7364
- **d**0

Solution:

A

Exercise:

Problem: What is the 40th percentile of wait times for customers calling the help line?

- **a** 6.87 minutes
- **b** 3.83 minutes
- **c** 0.68 minutes
- **d** 0.122 minutes

Solution:

B

Exercise:

Problem: The customer wait time that is 1 standard deviation above the mean is:

- **a** 2.17 minutes
- **b** 7.5 minutes
- **c** 9.67 minutes
- **d** 15 minutes

Solution:

D

Exercise:

Problem: The probability that a customer calling the help line waits exactly 6 minutes for help:

- **a** 0
- **b**0.45
- **c** 0.55
- **d**0.8

Solution:

A

Questions 33 – 35 refer to the following: ABC Delivery Service offers next day delivery of packages weighing between 2 and 20 pounds in a certain city. The have found that the weights of the packages they deliver are uniformly distributed between 2 and 20 pounds.

Exercise:

Problem: What is the probability that a package weighs between 10 and 15 pounds?

- **a** 0.2778
- **b** 0.5556
- **c** 0.2500
- **d**0.8333

Solution:

A

Exercise:

Problem:

Given that a package weighs less than 10 pounds, what is the probability that it weighs less than 5 pounds?

- **a** 0.1667
- **b** 0.6250
- **c** 0.3750
- **d**0.5000

Solution:

C

Exercise:

Problem: 35% of packages weigh less than how many pounds?

- **a** 7.8 pounds
- **b** 8.3 pounds
- **c** 11.7 pounds
- **d**13.7 pounds

Solution:

B

Exercise:

Problem:

Suppose that the probability that an adult in California will watch a Giant's World Series game is 65%. Each person is considered independent. Of interest, is the number of adults in California we must survey until we find one who will watch a Giant's World Series game. What is the probability that you must ask 2 or 3 people? (This is a geometric problem. If you did not cover the geometric distribution, then skip this problem.)

- **a** 0.6500
- **b** 0.3071
- **c** 0.2275
- **d**0.0796

Solution:

B

Questions 37– 39 refer to the following: The amount of time De Anza students work per week is approximately normally distributed with mean of 18.17 hours and a standard deviation of 12.92 hours.

Exercise:

Problem: The median is:

- **a** Not enough information
- **b** 12.92
- **c** 2.0
- **d** 18.17

Solution:

D

Exercise:

Problem: The 90th percentile for the amount of time De Anza students work per week is:

- **a** 1.61
- **b** 18.17
- **c** 90
- **d** 34.7

Solution:

D

Exercise:

Problem: Which of the following is NOT TRUE about the normal distribution?

- **a** the mean, median and mode are equal
- **b** the curve is skewed to the right
- **c** the curve never touches the x-axis

- **d** the area under the curve is one.

Solution:

B

Exercise:

Problem: We use the z-score to:

- **a** compare normal distributions with different averages and standard deviations
- **b** drive statistics students nuts
- **c** compare exponential distributions with the same average
- **d** compare uniform distributions with different minimum and maximum numbers

---

Solution:

A

Elementary Statistics: Skills Practice Exam 3- Lessons 7,8, 9, 10

Questions 1, 2, and 3 refer to the following:
In a study of a sample of 35 Computer Science majors and 40 Business majors, the average number of science fiction books each group read per year was recorded. The Computer Science majors read an average of 11 books with a standard deviation of 4 books. The Business majors read an average of 9 books with a standard deviation of 4.5 books. We are interested in whether the average number of science fiction books read by the Computer Science majors is the same as the number read by the Business majors. CS = Computer Science majors B = Business majors
Exercise:

Problem: The alternate hypothesis is:

- **a** mCS = mB
- **b**mCS ¹ mB
- **c** mCS < mB
- **d** mCS > mB

Solution:

B

Exercise:

Problem: The exact distribution for the test is:

- **a**Normal
- **b**Student-t with df » 72.9775
- **c**Exponential
- **d**Uniform

Solution:

C

Exercise:

Problem: The p-value is

- **a** 0.0452
- **b** 2.0376
- **c** 0.0226
- **d** 0

Solution:

A

Questions 4 and 5 refer to the following:
Suppose that a random survey of 10 teenagers found that the average amount of time they spend on the Internet each day is 3.2 hours with a sample standard deviation of 0.78 hours.

Exercise:

Problem:

The point estimate of the population average amount of time teenagers spend on the Internet each day is

- **a** 0.78 hours
- **b** 32 hours
- **c** 3.2 hours
- **d** unknown

Solution:

C

Exercise:

Problem:

An 87% confidence interval for the average amount of time a teenager spends on the Internet each day would be

- **a**(2.71, 3.68)
- **b** (2.83, 3.57)
- **c**(2.64, 3.76)
- **d**(2.79, 3.61)

Solution:

D

Exercise:

Problem:

Which of the following is TRUE about a 99% confidence interval for the true average amount of time that teenagers spend on the Internet each day?

- **I**We are 99% confident that the true average time that teenagers spend on the Internet per day lies within the confidence interval.
- **II**The confidence interval contains 99% of the data values collected.
- **III**99% of all the confidence intervals constructed this way contain the true average amount of time teenagers spend on the Internet each day.

- **a**I, II, and III
- **b**I and II
- **c**I and III
- **d**I only

Solution:

C

Questions 7 and 8 refer to the following:
Suppose a random survey of 600 registered voters revealed that only 39% actually voted in the last primary election. We are interested in the population proportion of registered voters who actually voted in the last primary election.

Exercise:

Problem:

The margin of error (error bound) for a 92% confidence interval for the true proportion of registered voters who actually voted in the last primary election is

- **a**0.0176
- **b**0.3900
- **c**0.3551
- **d**0.0349

Solution:

D

Exercise:

Problem:

Suppose you want to redo the survey of registered voters to see how many voted in the last election but this time you survey 1000 registered voters and find 340 voted in the last election. If the confidence level is maintained at 92% what happens to the confidence interval.

- **a**The confidence interval gets wider.
- **b**The confidence interval gets narrower.
- **c**The confidence interval stays the same.
- **d**The confidence interval cannot be calculated.

Solution:

B

Exercise:

Problem:

A study of a certain brand of AA batteries yielded a sample mean lifetime of 450 minutes with a sample standard deviation of 92 minutes. A hypothesis test was performed using the following hypotheses:

H_o : m = 480

H_a : m < 480

The type I error for this hypothesis test is:

- **a**to conclude that the average battery lifetime is less than 480 minutes when, in reality, it is equal to 480 minutes
- **b**to conclude that the average battery lifetime is equal to 480 minutes when, in reality, it actually is equal to 480 minutes
- **c**to conclude that the average battery lifetime is equal to 480 minutes when, in reality, it is less than 480 minutes
- **d**to conclude that the average battery lifetime is greater than 480 minutes when, in reality, it is equal to 480 minutes

---

Solution:

A

Exercise:

Problem:

The null hypothesis is “the percentage of men who score 45 points or more (out of 50 points) on a statistics test at De Anza College is the same as the percentage of women who score 45 points or more.” The alternate hypothesis is “the percentage of men who score 45 points or more (out of a 50 point test) on a statistics test at De Anza College is more than the percentage of women who score 45 points or more.”

The appropriate hypothesis test to perform and distribution to use is

- **a**a test of two population proportions, independent groups; Student-t distribution
- **b**a test of a single population proportion; Normal distribution
- **c**a test of two population means, independent groups; Normal distribution
- **d**a test of two population proportions, independent groups; Normal distribution

---

Solution:

D

Questions 11 through 13 refer to the following: In a study of vehicle safety, 15 minivans were crash tested and the repair costs for each of the 15 minivans were recorded. For these 15 minivans, the average repair cost was $1786 and the standard deviation was $937 (based on data from the Highway Loss Data Institute.) Suppose that you want to test the hypothesis that the average repair cost is under $2000. Assume that the underlying population of repair costs follows a normal distribution.
Exercise:

Problem: The correct null hypothesis for this test is

- **a**m ^{3} 1786
- **b**m £ 2000

- **c**m ^{3} 2000
- **d**m < 2000

Solution:

C

Exercise:

Problem:

At a 5% level of significance (a), the correct decision for this hypothesis test is

- **a**reject H_o because a is more than the p-value.
- **b**reject H_o because a is less than the p-value.
- **c**do not reject H_o because a is more than the p-value.
- **d**do not reject H_o because a is less than the p-value.

Solution:

D

Exercise:

Problem: The appropriate distribution for this test is

- **a**Normal with standard deviation $937
- **b**t with degrees of freedom = 15
- **c**t with degrees of freedom = 14
- **d**Normal with standard deviation $241.93

Solution:

C

For problems 14 and 15: Buses on a particular route stop in front of De Anza College every 20 minutes between 3:00 p.m. and 1:00 a.m. The waiting times are equally likely. We asked the 33 people waiting at 6:45 p.m. how long they had been waiting, and then calculated the average wait time for those people.

Exercise:

Problem: The distribution of the average wait times is:

- **a** N(10 , 1. 0050)
- **b** U(0 , 20)
- **c** N(10 , 5.7735)
- **d** Exp (1 20)

Solution:

A

Exercise:

Problem:

The probability that the average wait time is no more than 15 minutes is

- **a** 1
- **b** 0. 7500
- **c** 0. 7769
- **d** 0

Solution:

A

Exercise:

Problem:

Which probability statement best describes the graph? The horizontal axis has the label Xbar.

- **a** P(Xbar > 6)
- **b** P(Xbar < 3)
- **c** P(Xbar > 6 or X < 3)
- **d** P(3 < Xbar < 6)

Solution:

D

Questions 17 and 18 refer to the following: A radio news story claimed that half of all U.S. adults have Internet access. In a national poll about Internet usage (The Pew Internet Project), 12,638 U.S. adults were surveyed and it was found that 6413 of those surveyed had Internet access. At a 5% level of significance, perform a hypothesis test to test the claim made by the radio news story.

Exercise:

Problem: The hypothesis test is:

- **a** right - tailed
- **b** left - tailed
- **c** two – tailed
- **d** no-tailed

Solution:

C

Exercise:

Problem: The correct conclusion is:

- **a** The percentage of U.S. adults that have Internet access is not one-half.
- **b** more than half of all U.S. adults have Internet access.
- **c** half of all U.S. adults have Internet access.
- **d** less than half of all U.S. adults have Internet access.

Solution:

C

Questions 19 - 21 refer to the following: An organic fertilizer and a conventional chemical fertilizer are tested to determine if the organic fertilizer produces more blossoms per stalk. 43 pairs of seed (the pair of seeds come from the same parent plant) are treated. One seed is treated with the organic fertilizer and the other seed is treated with the conventional chemical fertilizer. After a growing season of identical watering and sunlight, the number of blossoms on each stalk in a matched pair is recorded. For each pair, the difference in the number of blossoms per stalk (organic fertilizer plant blossoms – conventional chemical fertilizer plant blossoms) is computed. _ The statistics are: xd = 2.2 sd = 5.3 n = 43
Exercise:

Problem: What type of hypothesis test is conducted?

- **a** Test of a single population mean
- **b** Test of two population proportions, independent groups
- **c** Matched or Paired Samples
- **d** Test of two population means, independent groups

Solution:

C

Exercise:

Problem: Select the appropriate alternate hypothesis.

- **a** µd £ 0
- **b** µd > 0
- **c** µd > 2.2
- **d**µd £ 2.2

Solution:

B

Exercise:

Problem: What is the correct conclusion?

- **a** The organic fertilizer does result in more blossoms per stalk
- **b** The organic fertilizer does not result in more blossoms per stalk.
- **c** The organic fertilizer results in the same number of blossoms per stalk
- **d** The conventional chemical fertilizer does result in more blossoms per stalk

---

Solution:

A

Problems 22 – 24 refer to the following: Based on data from the 2000 Census the average age of a Baldwin County, Alabama resident is 39 years with a standard deviation of 22. 6 . The data is normally distributed.

Exercise:

Problem: The median age, in years, is:

- **a** 39
- **b** 22.6
- **c** 19.5
- **d.** There is not enough information.

Solution:

A

Exercise:

Problem:

If 100 residents are surveyed, the IQR for the average age is:

- **a** 0. 30
- **b** 30.5
- **c** 3.04
- **d.**There is not enough information.

Solution:

C

Questions 24 – 25 refer to the following: Students doing a statistics project at Central City College found that the amount of time a teenager spends cleaning his/her room each week is exponentially distributed with a mean of 20 minutes.

Exercise:

Problem:

What is the probability that 50 randomly selected teenagers spend, on average, between 15 and 30 minutes cleaning their rooms each week?

- **a** 0.7500
- **b** 0.2902
- **c** 0.2492

- **d**0.9612

Solution:

D

Exercise:

Problem:

Find the 70th percentile for the average amount of time 50 teenagers spend cleaning their rooms each week.

- **a** 70
- **b** 21.5
- **c** 30.5
- **d**7.1

Solution:

B

Elementary Statistics: Exam 1: Lessons 1 & 2

Questions 1 – 2 refer to the following:
A sample of 56 employees was taken to determine average pulse rate. The data is in the table below:

Pulse Rate (beats per minute)	**Frequency (# of employees)**
54	1
58	4
65	7
68	9
72	10
76	4
80	10
84	6
90	3
98	2

Exercise:

Problem: The mode(s) is/are:

- **A**74
- **B**72 and 80
- **C**68
- **D**the average of 72 and 80

Solution:

B

Exercise:

Problem: The variable is:

- **A**the number of employees
- **B**all employees
- **C**the average pulse rate for the 56 employees
- **D**the pulse rate of one employee

Solution:

D

Questions 3 – 5 refer to the following:
The table below shows the ages of 50 senior citizens who attend a local senior center.

Age	Frequency	Relative Freq.	Cum. Rel. Freq.
56	9	0.18	

Age	Frequency	Relative Freq.	Cum. Rel. Freq.
63	8		
68	17	0.36	
73	8	0.16	
80	5	0.10	
88	2	0.04	1.00

Exercise:

Problem: Find the IQR.

- **A**0.5
- **B**17
- **C**10
- **D**not enough information

Solution:

C

Exercise:

Problem: Which interval has the smallest percentage of data?

- **A**The interval 55.5 to 63.5.
- **B**The interval 63.5 to 68.5.
- **C**The interval 68.5 to 73.5.
- **D**The interval 73.5 to 88.5.

Solution:

D

Exercise:

Problem: What percent of the ages are at most 68?

- **A**36
- **B**70
- **C**30
- **D**66

Solution:

B

Exercise:

Problem:

What can be said about a set of data when its standard deviation is zero?

- **A**If the data are ordered, they are very spread out from the mean.
- **B**All of the data appear with the same frequency.
- **C**If the data are ordered, they are very close to but different from the mean.
- **D**There is no mode.

Solution:

D

Exercise:

Problem:

For the following data, which measure of central tendency would be the LEAST useful? Data:

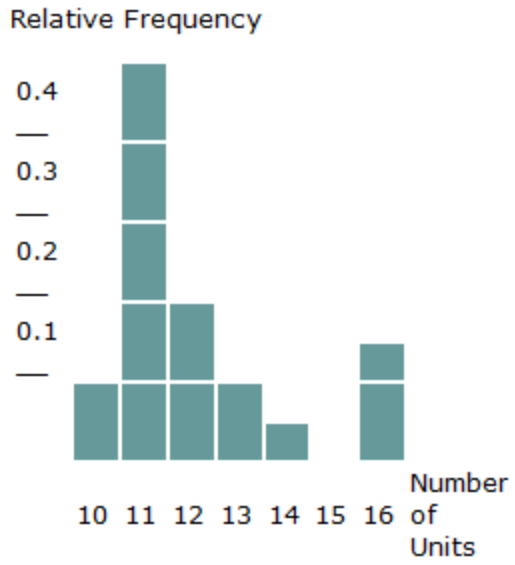
4 8 11 11 11 27 27 27 27 27 1000

- **A**mean
- **B**median
- **C**mode
- **D**standard deviation

Solution:

A

Questions 8 – 10 refer to the following:
Sixty (60) college students were asked the number of units they are taking this quarter. The results are given in the following graph:



Exercise:

Problem:

The number of responses that were “12” OR “13” is approximately:

- **A**18
- **B**0.3
- **C**25

- **D**Not enough information

Solution:

A

Exercise:

Problem: The third quartile is:

- **A**15
- **B**14
- **C**13
- **D**12

Solution:

C

# of Videos	Frequency	Relative Frequency	Cum. Rel. Freq.
0	18	0.29	?
1	26	0.41	?
2	?	?	?
3	6	?	?

# of Videos	Frequency	Relative Frequency	Cum. Rel. Freq.
4	1	?	?

Exercise:

Problem:

The sample from the table above was taken by randomly selecting one student from the administration's official list of students and then choosing every 100th. This is an example of what kind of sampling?

- **A**cluster
- **B**systematic
- **C**stratified
- **D**convenience

Solution:

B

Exercise:

Problem:

How should you classify data from the following question: What is your blood pressure?

- **A**qualitative
- **B**quantitative – discrete
- **C**quantitative – continuous

Solution:

C

Exercise:

Problem:

How should you classify data from the following question: What is your favorite vacation place?

- **A**qualitative
- **B**quantitative – discrete
- **C**quantitative – continuous

Solution:

A

Exercise:

Problem:

A study is done to determine the average amount of tuition all San Jose State undergraduate students pay per semester. A sample of 100 undergraduate San Jose State students is taken by dividing the students into freshmen, sophomore, junior, and senior years and selecting 25 students from each. Each student is asked how much tuition he/she paid for the Spring 2002 semester. The type of sampling and the parameter are:

- **A**cluster and the average amount of tuition all San Jose State students paid Spring 2002.
- **B**stratified and the average amount of tuition the 100 San Jose State students paid Spring 2002.
- **C**cluster and the average amount of tuition the 100 San Jose State students paid Spring 2002.
- **D**stratified and the average amount of tuition all San Jose State students paid Spring 2002.

Solution:

D

Elementary Statistics: Exam 2: Lessons 3 & 4

Questions 1 – 4 refer to the following:
The following table classified eighty-two children by age and favorite meals at a McDonald's ® restaurant.

	6 or below	**7-12**	**13-15**	**Total**
Hamburger	6	12	16	34
Chicken Nuggets	9	13	11	33
Filet ‘o’ fish	2	5	8	15
Totals	17	30	35	82

Age (in years) against favorite foods

Assume a child is randomly selected.
Exercise:

Problem:

Find the probability of being “7-12” years old AND preferring “chicken nuggets”

- **A** $\frac{13}{82}$
- **B** $\frac{28}{82}$
- **C** $\frac{33}{82}$
- **D** $\frac{13}{61}$

Solution:

A

Exercise:

Problem:

Find the probability of being “13-15” years old OR preferring “Filet’o’fish”.

- **A** $\frac{8}{82}$
- **B** $\frac{50}{82}$
- **C** $\frac{8}{51}$
- **D** $\frac{42}{82}$

Solution:

D

Exercise:

Problem:

Find the probability of “preferring Hamburger” given that the randomly selected child is 13-15 years old.

- **A** $\frac{16}{82}$
- **B** $\frac{16}{35}$
- **C** $\frac{16}{34}$
- **D** $\frac{16}{70}$

Solution:

B

Exercise:

Problem:

The events “preferring Hamburger” and “being 13-15 years old” are:

- **A**Mutually exclusive
- **B**Independent
- **C**Neither mutually exclusive or independent.
- **D**Both mutually exclusive and independent.

Solution:

C

Exercise:

Problem:

E and F are two events such that $P(E) = 0.60$, $P(E \text{ or } F) = 0.90$ and $P(E \text{ and } F) = 0.50$. Find $P(F)$.

- **A**0.80
- **B**0.30
- **C**0.40
- **D**0.10

Solution:

A

Exercise:

Problem:

The probability that a randomly chosen adult resident of Bayview city owns a boat is 0.16. The probability that a randomly chosen adult rents an apartment is 0.30. The probability that the adult owns a boat given he/she rents an apartment is 0.20.

- **A** 0.048
- **B** 0.24
- **C** 0.10
- **D** 0.06

Solution:

D

Exercise:

Problem: Possessing a boat and renting an apartment are:

- **A** independent events
- **B** mutually exclusive
- **C** both independent and mutually exclusive
- **D** neither independent nor mutually exclusive

Solution:

D

Questions 8 – 9 refer to the following:
A bag contains 4 red marbles and 5 blue marbles. Two marbles are randomly drawn without replacement.

Exercise:

Problem:

Find the probability of the event "The first marble is red and the second is blue."

- **A** $\frac{20}{81}$
- **B** $\frac{20}{72}$
- **C** $\frac{4}{12}$
- **D** $\frac{4}{9}$

Solution:

B

Exercise:

Problem: Find the probability that both marbles are red.

- **A** $\frac{16}{81}$
- **B** $\frac{7}{81}$
- **C** $\frac{12}{72}$
- **D** $\frac{8}{72}$

Solution:

C

Exercise:

Problem:

Approximately 70% of U. S. adults had at least one pet as a child. We randomly survey 60 U. S. adults. We are interested in the number that had at least one pet as a child. The probability that at least 3 adults had at least one pet as a child means:

- **A** $P(X = 0) + P(X = 1) + P(X = 2)$
- **B** $P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$
- **C** $P(X = 4) + P(X = 5) + P(X = 6) + ... + P(X = 60)$
- **D** $P(X = 3) + P(X = 4) + P(X = 5) + ... + P(X = 60)$

Solution:

D

Questions 11 – 12 refer to the following:
A plumber has determined the possible number of house calls to be made each day, and their related probabilities:

$x =$**# of house calls**	$P(x)$
0	0.10
1	0.40
2	0.25
3	0.15
4	0.10

Exercise:

Problem:

What is the probability that he makes at least 1, but no more than 3 house calls in a day?

- **A**0.65
- **B**0.80
- **C**0.50
- **D**0.40

Solution:

B

Exercise:

Problem:

If the plumber charges a flat fee of $ 40 for a house call, the expected daily income is:

- **A**$70
- **B**$175
- **C**$400
- **D**$1.75

Solution:

A

Exercise:

Problem: If X~$B(40,0.2)$, then $P(X > 11) =$

- **A**0.0432
- **B**0.0001
- **C**0.0875
- **D**0.1608

Solution:

C

Exercise:

Problem:

The Fizz–Full Soda Company knows that 4% of the bottles of soda it produces are filled with less soda than required. If one purchases 10 bottles at random, the probability that at most 2 of these bottles will have less soda than required is:

- **A**0.0519
- **B**0.9938
- **C**0.9418
- **D**0.0080

Solution:

B

Questions 15 – 16 refer to the following:
Assume the statistics final is a multiple-choice exam with 50 questions, each question having 5 choices, only one of which is correct. Assume you answer all questions at random (guessing).
Exercise:

Problem:

The expected number of questions you would get correct would be:

- **A**5
- **B**10
- **C**40
- **D**45

Solution:

B

Exercise:

Problem:

Based upon numerical calculations, would you be surprised if a person got exactly half of the questions correct?

- **A**yes, because it is impossible
- **B**yes, because the probability is almost 0
- **C**no, because the probability is 0.50

- **D**no, because it is the most likely probability

Solution:

B

Exercise:

Problem:

If sampling without replacement occurs, do the picks follow the Binomial Distribution?

- **A**Yes, because each pick is independent from the others.
- **B**No, because the probability of success on each pick changes.
- **C**Yes, if we are counting the number of successes.
- **D**No, because we may not have any successes.

Solution:

B

Exercise:

Problem:

Ninety-four percent of California community college transfers feel that their community college adequately prepared them to handle upper-division coursework at their transfer university. We randomly survey 14 California community college transfers. We are interested in the number that feel that their community college adequately prepared them to handle upper division coursework at their transfer university. List the values that X , the Random Variable, may take on.

- **A**$1,2,3,...,14$
- **B**$1,2,3,...,94$
- **C**$0,1,2,...,14$
- **D**$0,1,2,...,94$

Solution:

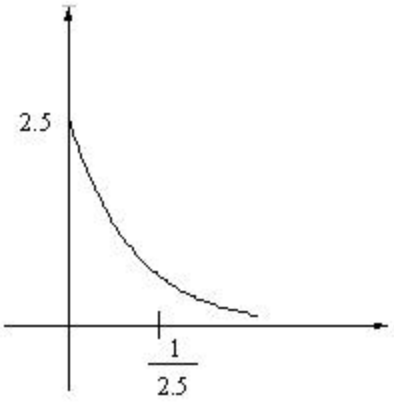
C

Elementary Statistics: Exam 3: Lessons 5, 6 & 7

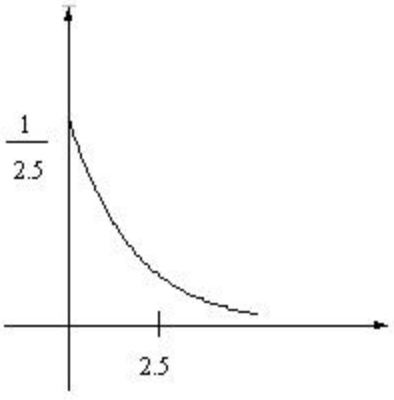
Questions 1 – 3 refer to the following:
Assume the amount of money seventh–grade students spend on food each day at school is exponentially distributed with an average of $2.50.
Exercise:

Problem: Which graph best describes the distribution?

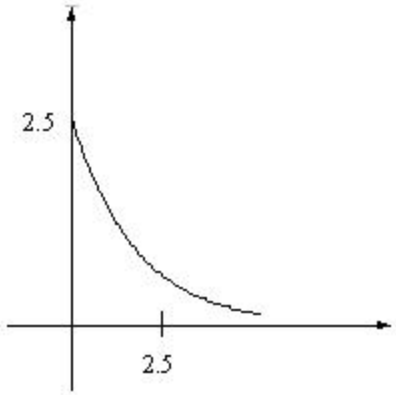
- **A**



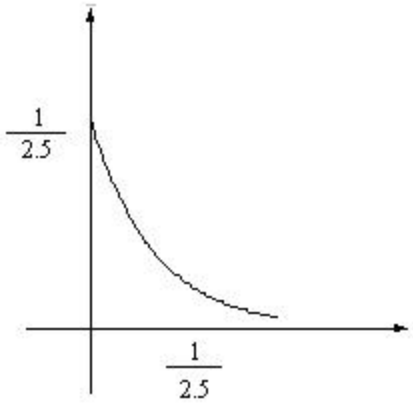
- **B**



- **C**



- **D**



Solution:

B

Exercise:

Problem:

Find the probability that a randomly selected seventh–grade student spends less than $4 a day on food.

- **A** 0.7981
- **B** 0.2019
- **C** 0.9999
- **D** 0.0001

Solution:

A

Exercise:

Problem:

85% of the seventh–grade students spend more than what amount per day?

- **A** $2.12
- **B** $0.75
- **C** $4.74
- **D** $0.41

Solution:

D

Questions 4 – 5 refer to the following:
The amount of time that intermediate algebra students at Leland High School spend doing their homework per day is normally distributed with a mean 1.5 hours and standard deviation 0.75 hours.

Exercise:

Problem:

If one student is randomly chosen, what is the probability that the student does intermediate algebra homework at least 2 hours per day?

- **A** 0.7475
- **B** 0.4259
- **C** 0.2525
- **D** 0.6784

Solution:

C

Exercise:

Problem:

60% of these students spend at most how many hours doing their homework?

- **A**1.69 hours
- **B**1.31 hours
- **C**1.5 hours
- **D**0.2533 hours

Solution:

A

Questions 6 – 7 refer to the following:
Llamas are excellent pack animals. It is known that the weight of supplies carried by llamas follows a normal distribution with a mean of 62.5 pounds and a standard deviation of 6 pounds.

Exercise:

Problem:

Find the probability that the weight of supplies carried by one randomly chosen llama is between 60 and 70 pounds.

- **A**0.4441
- **B**0.5559
- **C**0.8944
- **D**1

Solution:

B

Exercise:

Problem:

The middle 50% of weights of supplies carried by a randomly chosen llama is between _____ and _____.

- **A**0 and 62.5 pounds
- **B**58.45 and 66.55 pounds
- **C**56.5 and 68.5 pounds
- **D**There is not enough information given.

Solution:

B

Exercise:

Problem: Which of the following are true for the normal distribution?

- **I**More values fall close to the mean than fall far away from the mean.
- **II**The mean and standard deviation cannot be the same.
- **III**A change in μ causes the graph to shift to the left or right and changes the shape of the graph.
- **IV**A change in s causes a change in the shape of the normal curve.

- **A**I, IV
- **B**I, II, III, IV
- **C**I, II, III
- **D**III, IV

Solution:

A

Questions 9 – 13 refer to the following:
The length of time junior high school students sleep per night follows an approximate uniform distribution from seven to eleven hours. Suppose we randomly select a junior high student.

Exercise:

Problem:

Find the probability that the randomly selected student sleeps less than 8 1/2 hours per night.

- **A**.2143
- **B**0.7727
- **C**0.4705
- **D**0.375

Solution:

D

Exercise:

Problem:

Find the probability that the randomly selected student sleeps eight to twelve hours per night.

- **A**0
- **B**1
- **C**0.75
- **D**0.25

Solution:

C

Exercise:

Problem:

On average, how long does a junior high school student sleep per night?

- **A**.2143
- **B**0.7727
- **C**0.4705
- **D**0.375

Solution:

B

Exercise:

Problem:

On average, how long does a junior high school student sleep per night?

- **A**9.6 hours
- **B**6.5 hours
- **C**7.8 hours
- **D**8.4 hours

Solution:

D

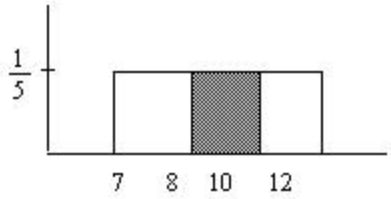
Exercise:

Problem:

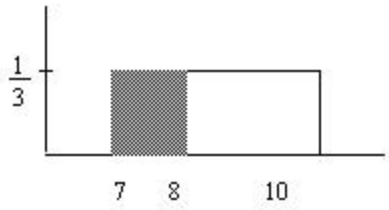
We are interested in the probability that a randomly selected student sleeps less than eight hours, knowing that he/she sleeps less than ten. Which graph best depicts this situation?

- **A**

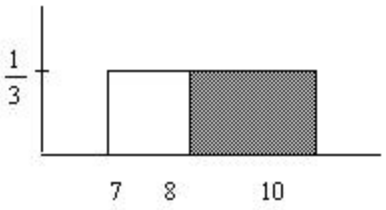
- **B**



- **C**



- **D**



Solution:

C

Elementary Statistics: Exam 4: Lessons 8, 9 & 10

- Each question has exactly one best answer.
- Each question is worth 3 points for a total of 54 points (4 bonus points possible).

Questions 1 – 3 refer to the following:
Last year, President Bush granted a tax-cut check to all income tax filers. In doing so, it was reported that he thought that at least 30% of the households would use the tax-cut check to increase spending. According to a report by the University of Michigan Research Center (Wall Street Journal, Feb. 26, 2002), 220 of the 1000 people surveyed said that the 2001 tax-cut check they received has led them to increase spending.

Exercise:

Problem: The alternate hypothesis for this test is:

- **A** $p>0.30$
- **B** $p>0.22$
- **C** $p>0.30$
- **D** $p>0.22$

Solution:

C

Exercise:

Problem: A Type II error is to:

- **A** Conclude that the proportion of people that would use the tax-cut check to increase spending is less than 30% when in fact the proportion is at least 30%.
- **B** Conclude that the proportion of people that would use the tax-cut check to increase spending is at least 30% when in fact the proportion is less than 30%.
- **C** Conclude that the proportion of people that would use the tax-cut check to increase spending is less than 22% when in fact the proportion is at least 22%.
- **D** Conclude that the proportion of people that would use the tax-cut check to increase spending is at least 22% when in fact the proportion is less than 22%.

Solution:

B

Exercise:

Problem: Which of the following is the correct decision to make for the test?

- **A** Reject the null hypothesis.
- **B** Do not reject the null hypothesis.
- **C** The test is inconclusive.

Solution:

C

Questions 4 – 6 refer to the following:
According to an article by George Will (San Jose Mercury News, Feb. 28, 2002), the average U.S. consumption per person per year of French Fries is 28 pounds. Suppose that you believe that the average in Santa Clara County is not 28 pounds. You randomly survey 50 people in this county. The sample average is 24 pounds with a sample standard deviation of 10 pounds. Conduct an appropriate hypothesis test.
Exercise:

Problem: This test is:

- **A**left-tailed
- **B**right-tailed
- **C**two-tailed
- **D**no-tailed

Solution:

C

Exercise:

Problem: The p-value for this test is:

- **A**0.0068
- **B**0.0034
- **C**0.0047
- **D**0.0136

Solution:

A

Exercise:

Problem: At the 5% level, the correct conclusion is:

- **A**The average consumption in Santa Clara County is less than 28 pounds.
- **B**The average consumption in Santa Clara County is not 28 pounds.
- **C**The average consumption in Santa Clara County is less than 24 pounds.
- **D**The average consumption in Santa Clara County is 24 pounds.

Solution:

B

Questions 7 and 8 refer to the following:
A hospital administrator wants to determine the proportion of emergency room patients that use the emergency room (ER) for non-emergency care. She randomly samples records from 350 ER patients and determines that 74 of those patients required only non-emergency care.

Exercise:

Problem:

The administrator constructs a 90% confidence interval for the true proportion of all ER patients who use the ER for non-emergency care. The error bound for the proportion EBP for this confidence interval is:

- **A**0.036
- **B**0.072
- **C**0.030
- **D**0.106

Solution:

A

Exercise:

Problem:

If the same data were used but the confidence level used was 95% instead of 90%, the error bound for the proportion EBP would be:

- **A**larger
- **B**the same
- **C**smaller
- **D**we are unable to determine unless another sample is obtained

---

Solution:

A

Exercise:

Problem:

What is meant by the term “95% confident” when constructing a confidence interval for a mean?

- **A**If we took repeated samples, approximately 95% of the samples would produce the same confidence interval.
- **B**If we took repeated samples, approximately 95% of the confidence intervals calculated from those samples would contain the sample mean.
- **C**If we took repeated samples, the sample mean would equal the population mean in approximately 95% of the samples.
- **D**If we took repeated samples, approximately 95% of the confidence intervals calculated from those samples would contain the true value of the population mean.

Solution:

D

Questions 10 through 11 refer to the following:
An aircraft manufacturer is testing a new procedure to be used in installing a certain component in an aircraft. For a random sample of 8 airplanes being assembled, the time (in minutes) required to install the component for each of these 8 aircrafts are:

808487919195102106

Assume the underlying population of installation times is approximately normally distributed.
Exercise:

Problem:

Find the 90% confidence interval for the true mean installation time using this new procedure.

- **A**(86.1, 97.9)
- **B**(86.9, 91.1)
- **C**(87.2, 96.8)
- **D**Not enough information

Solution:

A

Exercise:

Problem: The value that is the center of the confidence interval is:

- **A**91
- **B**92
- **C**93
- **D**μ

Solution:

B

Exercise:

Problem:

The amount of soda contained in a can for a certain brand of soda is normally distributed with a population standard deviation of 0.1 ounces. A random sample of 40 cans of soda was selected and the amount of soda in each can was measured. The sample mean was 12.03 ounces and the sample standard deviation was 0.08 ounces.

What is the appropriate distribution to use when calculating a confidence interval for the true mean amount of soda contained in all cans of this brand?

- **A**The student-t distribution, because the sample standard deviation is given.
- **B**The student-t distribution, because the repair costs are approximately normally distributed.
- **C**The standard normal distribution, because the population standard deviation is known.
- **D**The standard normal distribution, because the sample mean is known.

Solution:

C

Questions 13 – 15 refer to the following:
Two competing parcel-delivery firms in a large city make conflicting claims about which one delivers parcels in the shortest time. A random sample of 100 delivery times for the first company produces a sample mean of x1 = 37 minutes and a sample standard deviation of s1 = 10 minutes. A random sample of 100 delivery times for the second company produces a sample mean of x2 = 41 minutes and a sample standard deviation of s2 = 12 minutes. Conduct a hypothesis test to determine if the mean delivery time for the first company is less than that for the second company.
Exercise:

Problem:

What is the appropriate distribution to use when calculating a confidence interval for the true mean amount of soda contained in all cans of this brand?

- **A**m1 < m2
- **B**m1 > m2
- **C**m1 £ m2
- **D**m1 ³ m2

Solution:

D

Exercise:

Problem: The exact distribution for the hypothesis test is:

- **A**The normal distribution because the population standard deviations are given.
- **B**The student-t distribution because the population standard deviations are unknown.
- **C**The exponential distribution because the time of delivery decreases.
- **D**Not able to determine.

Solution:

B

Exercise:

Problem: f the p-value is 0.0056, the conclusion is:

- **A**The mean delivery time for the first company is higher than the mean delivery time for the second company.
- **B**The mean delivery time for the first company is no more than the mean delivery time for the second company.
- **C**The mean delivery time for the first company is at least equal to the mean delivery time for the second company.
- **D**The mean delivery time for the first company is less than the mean delivery time for the second company.

Solution:

D

Questions 16 – 17 refer to the following:
Participants in a random sample of 10 professional football players are placed on a yogurt-and-banana diet for one month. The weights before and after one month on the diet are as follow:

Before	187	205	165	193	199	286	212	189	242	253
After	175	193	167	190	197	240	210	189	221	255

We want to determine if the yogurt-and-banana diet helps reduce the weight of the football players. Assume the weight of the professional football players is approximately normally distributed.
Exercise:

Problem: This is a test of:

- **A**Two independent population means, population standard deviations known.
- **B**Two independent population means, population standard deviations unknown.
- **C**Paired or matched samples.
- **D**Two population proportions.

Solution:

C

Exercise:

Problem: The distribution for the hypothesis test is:

- **A**Student-t
- **B**Exponential
- **C**Normal

- **D**Uniform

Solution:

A

Exercise:

Problem:

A newspaper/TV network survey was conducted to determine whether the percentage of adult males who favor the death penalty is greater than the percentage of adult females who favor the death penalty. A random sample of 800 adult males produced 480 who favor the death penalty., while a random sample of 800 adult females produced 410 who favor the death penalty.

The Type I error is to:

- **A**Conclude that the percentage of adult males who favor the death penalty is greater than the percentage of adult females who favor the death penalty when, in fact, the percentage of males is no more than the percentage of females.
- **B**Conclude that the percentage of adult males who favor the death penalty is no more than the percentage of adult females who favor the death penalty when, in fact, the percentage of males is more than the percentage of females.
- **C**Conclude that the percentage of adult males who favor the death penalty is greater than the percentage of adult females who favor the death penalty when, in fact, the percentage of males is actually greater than the percentage of females.
- **D**Conclude that the percentage of adult males who favor the death penalty is no more than the percentage of adult females who favor the death penalty when, in fact, the percentage of males is actually no more than the percentage of females.

Solution:

A

Elementary Statistics: Exam 5: Lesson 11

Exercise:

Problem:

Suppose the random variable X follows a chi-square distribution with degrees of freedom equal to 35. Fill in the blanks.

- **A** $m =$
- **B** $s =$

Solution:

A. 35

B. 8.3666

Exercise:

Problem: Check all that apply.

- **A** The chi-square graph always has the same shape.
- **B** If X follows a chi-square distribution with df = 200, then X approximately follows a normal distribution.
- **C** The chi-square distribution is skewed to the right if the degrees of freedom are less than 90.
- **D** The test statistic for the chi-square distribution may be zero.
- **E** A goodness-of-fit hypothesis test is always right-tailed.
- **F** A test of independence tests whether two factors are independent or not.

Solution:

B, C, D, and F

Exercise:

Problem: Write the null and alternate hypotheses for the following:

It is believed that public high school students attend school in equal numbers for each day of the school week. Suppose a sample of the days students were present at school was taken for one particular high school:

- 1750 students were present on Monday
- 1800 students were present on Tuesday
- 1840 students were present on Wednesday
- 1810 students were present on Thursday
- 1800 students were present on Friday

Ho:

Ha:

Solution:

Ho: Public high school students attend school in equal numbers for each day of the school week

Ha: Public high school students DO NOT attend school in equal numbers for each day of the school week

Exam 6: Chapter 12
Practice final exam for use with Collaborative Statistics (col10522) by Barbara Illowsky and Susan Dean.

Questions 1 – 3 refer to the following:
New snowboarders (those who have snowboarded a year or less) often suffer from minor injuries. A random sample of seven new snowboarders produced the data on number of months snowboarding and number of minor injuries in the last month that they snowboarded.

Months Snowboarding	# of injuries in the last month
2	9
4	6
8	3
12	2
1	9
5	6
9	2

Exercise:

Problem:

Is the correlation between the number of months snowboarding and the number of injuries in the last month snowboarding significant?

- **A**Yes
- **B**No
- **C**Not enough information to answer question

Solution:

A

Exercise:

Problem: The linear regression equation is:

- **A**yhat = 9.5904 – 0.9614 x
- **B**yhat = -0.7349 + 9.5904x
- **C**yhat = 9.5904 + 0.7349x
- **D**yhat = 9.5904 – 0.7349x

Solution:

D

Exercise:

Problem:

If a new snowboarder has snowboarded for five (5) months, how many injuries would s/he have in the last month snowboarding?

- **A**4.8
- **B**47.2
- **C**5.9
- **D**13.3

Solution:

C

Exercise:

Problem:

If you calculate the line of best fit and the independent variable and the dependent variable have negative correlation, then the line of best fit has slope zero (0).

- **A**True
- **B**False
- **C**There is not enough information

Solution:

B

Questions 5 – 6 refer to the following:
The cost of a leading gourmet ice cream in different sizes is given in the table.

Size (ounces)	Cost
16	$4.29
32	$7.36
64	$12.80
96	$17.28

Exercise:

Problem: Are there any outliers?

- **A**Yes
- **B**No
- **C**Not enough information

Solution:

B

Exercise:

Problem:

If your friend used the line of best fit to predict the cost for a 128-ounce size of gourmet ice cream, what would you tell him/her with what you have learned about linear regression?

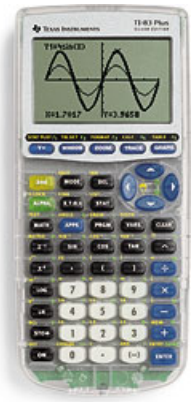
- **A**That a line is not the best fit for the cost of the gourmet ice cream.
- **B**That the 128-ounce size of the gourmet ice cream is far too expensive.
- **C**That we should switch the independent and dependent variables.
- **D**That s/he should not use the line of best fit to make this prediction.

Solution:

D

Elementary Statistics: Calculator Instructions and Resources

Graphing Calculator Instructions



- TI-83 video tutorials are available here: (Part 1, Part 2).
- The Discrete Mean & Stdev link shows how to calculate a mean and stdev for a discrete random variable.
- The Outliers link shows how to calculate outliers using the lists.
- The Matrix: TI-89 link shows how to create a matrix using the TI-89 calculator for use in a **Test of Independence**.
- The Random Numbers link shows how to generate random numbers.
- The Sorting link shows how to sort a list.
- The TI-89 link gives some general instructions for the TI-89 plus linear regression instructions.
- Link to Graphing Calculator Help for Several TI Calculators

Calculating the Mean and Standard Deviation of a Discrete Probability Distribution on the TI-83 and TI-86

1. Enter the possible values for your random variable X in list `L1`.
2. Enter the probabilities for each value of X in list `L2` in the position next to the value of X.
3. Calculate "One Variable Statistics" for lists `L1` and `L2`.
4. (Do this the same way you calculate one variable statistics for data values and frequencies when given a set of data.)

5. The calculator will give the mean as "x-bar" $\bar{x}$ although we know that this mean is actually μ because it is the mean of a probability distribution.
6. The calculator will give the standard deviation as σ which indicates the standard deviation of a probability distribution (as well as the standard deviation of a population).
7. The calculator does not give you a value for s because the frequencies you gave it are not whole numbers.

Outlier Instructions for the TI-83, 86, and 89 Calculators

When you finish going over these instructions, do TEXT problem #5 ("stories" and "height of building") in Ch. 12. One of the points is an outlier.

This explains how to find outliers on various calculators. Suppose the data is (3,5), (6,8), (9,7), (5,20). The `xlist` is 3, 6, 9, 5 and the `ylist` is 5, 8, 7, 20. Put the `xlist` into `L1` and the `ylist` into `L2`. Do the linear regression. `yhat = 11.2267 - .2133x`

Note:TI-89: Use Flashapps to enter your lists and then find your instructions below the instructions for the TI-83 & 86.

For TI-83 & 86 Calculators

1. Go back to where you entered the lists `L1` and `L2` and go to the list name `L3`. Enter `11.2267-.2133L1` (You enter the equation at the bottom of the screen where it says `L3`. Press `Enter`. In `L3` are the `yhat` values.
2. Arrow to the list name `L4`. Enter `L2-L3`. Press `Enter`. (In `L4` are the `y - yhat` values.)
3. Arrow to the list name `L5`. Enter `L4^2`. Press `Enter`. (In `L5` are the `(y - yhat)^2` values.)

4. Continue by finding the Instructions for the appropriate calculator below (83 or 86):

For TI-83

1. Exit to the Home Screen. Clear it. Press `2nd LIST`. Arrow to `MATH`. Press `5:sum`. (Press `L5`). Press `Enter`. You should see 137.1467 (to 4 decimal places). This is the SSE.
2. Calculate s. Press the square root symbol and enter `(137.1467/2)`. (You get the denominator by taking the number of data points and subtracting 2: $4 - 2 = 2$.) Press `Enter`. You should see 8.2809 (to 4 decimal places).
3. Multiply 8.2809 by 1.9. You should see 15.7337.
4. Press `L4`. Press `Enter`. (Use the arrow keys to scroll through the list.)
5. Compare 15.7337 to the absolute values of the numbers in `L4`. If any absolute value is greater than or equal to 15.7337, then the corresponding point is an outlier.
6. Absolute values of the numbers in `L4` are 5.587, 1.947, 2.307, 9.8398. None of them are greater than 15.7337, so no point is an outlier.

For TI-86

1. Exit to the Home Screen. Clear it. Press `2nd LIST`. Press `F5 (OPS)`. Press `MORE`. Press `F1 (sum)`. Press `2nd F3 (NAMES)`. Find the `F key` with `L5` and press it. Press `Enter`. You should see 137.1467 (to 4 decimal places). This is the SSE.
2. Calculate s. Press the square root symbol and enter `(137.1467/2)`. (You get the denominator by taking the number of data points and subtracting 2: $4 - 2 = 2$.) Press `Enter`. You should see 8.2809 (to 4 decimal places).
3. Multiply 8.2809 by 1.9. You should see 15.7337.
4. Press `L4`. Press `Enter`. (Use the arrow keys to scroll through `L4`.)
5. Compare 15.7337 to the absolute values of the numbers in `L4`. If any absolute value is greater than or equal to 15.7337, then the

corresponding point is an outlier.
6. Absolute values of the numbers in `L4` are 5.587, 1.947, 2.307, 9.8398. None of them are greater than 15.7337 so no point is an outlier.

For TI-89

1. **NOTE:** When you do Linear Regression, a list called "`resid`" is created automatically. This list has the `y - yhat` values in it.
2. Go into `Flashapps` and into your lists. Arrow to "`resid`" so you see it. Then arrow to the list name `L3`. Press `F3`. Press `1:names`. Arrow down to `STATVARS`. If it is not checked, press `F4`. Then arrow down to `resid`. Press `Enter`. Press `^2`. `L3` will have the `(y - yhat)^2` values. Press `HOME` and Press `F1 8` to clear it. Press `2nd MATH`. Press `3:List`. Press `6:Sum`. Press `alpha L3`. Depending on how you have `MODE Display Digits` set, you should see approximately 137.1467. This is the SSE.
3. Calculate s. (You should be `HOME`.) Press `clear`. Press the square root symbol and enter `137.1467/2`). (You get the denominator by taking the number of data points and subtracting 2: $4 - 2 = 2$.) Press `Enter`. You should see 8.2809 (to 4 decimal places).
4. Press the times key and enter `1.9`. Press `Enter`. You should see 15.7337
5. Press `clear`. Press `2nd VAR-LINK`. Arrow down to `resid` (It is below `STAT VARS`). Press `Enter`. Arrow up to the list. Scroll through the list using the arrow keys.
6. Compare 15.7337 to the absolute values of the numbers in the list. If any absolute value is greater than or equal to 15.7337, then the corresponding point is an outlier.
7. Absolute values of the numbers in the list are approximately 5.59, 1.95, 2.31, 9.84. None of them are greater than 15.7337 so no point is an outlier.

Create a Matrix to use in a Test of Independence on the TI-89

1. Press `APPS 6:Data/Matrix Editor`

2. Press `3:New`
3. Arrow over and down to `2:Matrix`.
4. Press `Enter`
5. Arrow down to `Folder`. Either use the one that is there or arrow over and down to another folder name (don't use `statvars`) and press `Enter`.
6. Arrow down to `Variable` and enter a name you will remember.
7. Arrow down to `Row dimension` and enter the number of rows you want.
8. Arrow down to `Column dimension` and enter the number of columns you want. (**NOTE:** You can change these numbers if you want for a different problem.)
9. Press `Enter` until you see your matrix with zeroes as the entries. Fill in your matrix with the data from the table.
10. Press `APPS`
11. Press `1:Flashapps`
12. Press `Enter`
13. Press `F6 TESTS`
14. Press `8:Chi2 2-way`
15. Enter the name of your matrix at `Observed Mat:`
16. Press `Enter`.
17. You should see the screen with the test statistic and the p-value.

- To change the size of your matrix, press `APPS 6:Matrix/Data Editor 2:Open`.
- Then fill in the `OPEN` screen with `Matrix`, the correct folder, and the correct name of your matrix.
- Press `Enter` until you see your matrix.
- To resize it, press `F6 Util 6:Resize Matrix`. Enter the row dimension and arrow down to column dimension and enter that number. Press `Enter` until you see your resized matrix.

Generating Random Numbers with the TI Calculators

The following functions will generate random numbers from the list of numbers 1 - 60.

TI-83

1. Press `MATH`. Arrow over to `PRB`.
2. Press `5:randInt`. Enter `(1,60)`.
3. Press `Enter` and you will see the first random number.
4. Keep pressing `Enter` to get random numbers between 1 and 60, inclusive.

TI-86

1. Press `2nd MATH`.
2. Press `PROB` (in `F2`).
3. Press `randint` (in `F5`). Enter `(1,60)`.
4. Press `Enter` and you will see the first random number.
5. Press `Enter` and you will see the first random number.

TI-89

- Press `2nd MATH`.
- Press `7:Probability`.
- Press `4:rand`. Enter `(1,60)`and press `Enter`.
- You will see the first random number.
- Keep pressing `Enter` to get random numbers between 1 and 60, inclusive.

How to Sort Data Using the TI Calculators

Only sort data where the frequencies are all 1.

TI-83

1. Enter the data into a list `L1`, `L2`, `L3`, `L4`, `L5`, or `L6`.

2. Press `2nd LIST` (above the `STAT` key).
3. Arrow over (use the right arrow) to `OPS`.
4. Press `1` (for `1:SORTA`).
5. Enter the name of your list and a right parenthesis. Press `Enter`.
6. Press `STAT 1` and look at your list. It should be in sorted order.

TI-86

1. Enter the data into a list.
2. Press `EXIT`.
3. Press `2nd LIST` (above – key).
4. Press `OPS` (in `F5`). Press `sortA` (in `F2`).
5. Enter the list name by pressing `2nd NAMES` (in `M3`).
6. Arrow to your list and press the appropriate `F` key.
7. Press the `STO->` key.
8. Enter the same list name. Press `Enter`.
9. Check that your list is sorted by going back into `2nd STAT EDIT`.

TI-89

> **Note:**For the TI-89, get Flashapps from the TI Web site or have your instructor put it into your calculator.

1. Press `APPS 1` (for `1:Flashapps`). Press `Enter`.
2. Press `Enter` again. Enter your data into a list (Enter data into one of the L lists, `L1`, `L2`, `L3`, `L4`).
3. Press `F3` (for `List`). Press `2` (for `2:Ops`).
4. Press `1` (for `1:Sort List`). If you see your list name, press `Enter`. If the correct list name is not there, then enter it. Press `Enter`.
5. You should see your list in sorted order.

Notes for the TI-89 Calculator

In most TI-89 calculators, you must load a new operating system first and then the APPS statistics program. You can either go to your instructor OR you can go to the TI-89 site to get the operating system and the APPS statistics program. You must go to the TI-89 site to get the Guidebook. It has all the instructions for running the statistics program.

To go to the TI-89 site, click http://education.ti.com to get to the TI-89 link. Then choose "apps" and "Handheld Software Application." From this page, you can download the operating system (version 2.05) and statistics program (Statistics with List Editor - this page will have both the operating system and the program).

You must then load the Advanced Mathematics Software Operating System and the Statistics with List Editor from your computer to the TI-89 calculator (unless your instructor has loaded them directly into your calculator). For that, you will need the TI-GRAPH LINK™ software and the TI-GRAPH LINK™ Cable. TI sells them on their Web Site.

Getting Started

Create the folder "`mystat`" (for statistics):

1. Press `2nd VAR-LINK`. Press `F1 Manage`.
2. Press `5:Create Folder`. Enter the name "`mystat`" and press `Enter` twice.

Make `mystat` your current folder:

1. Press `MODE`. Arrow down to `Current Folder` and arrow over and down to `mystat`.
2. Press `Enter` twice. You should see the name `MYSTAT` in the lower left corner.

Creating list names `L1`, `L2`, and `L3`:

1. Press `APPS 1:Flashapps`. Press `Enter`.
2. Arrow up into the name area of the lists and over until you reach a blank title area.
3. Press `L` (above the `4`). Press `alpha`. Press `1`. Press `Enter`.
4. Arrow up into the name area of the lists and over until you reach a blank title area.
5. Press `L`. Press `alpha`. Press `2`. Press `Enter`.
6. Arrow up into the name area of the lists and over until you reach a blank title area.
7. Press `L`. Press `alpha`. Press `3`. Press `Enter`.

Linear Regression

Given the data set: (3, 5), (6, 8), (9,7), (5, 20). The `xlist` is 3, 6, 9, 5 and the `ylist` is 5, 8, 7, 20. Put the `xlist` into `L1` and the `ylist` into `L2`.

Constructing a Scatter Plot

1. Make sure you are using Flashapps (Press `APPS 1:Flashapps`. Press `Enter`.)
2. **SCATTERPLOT:** After you are have entered your lists in `L1` and `L2`, press `F2`.
3. Press `1:Plot Setup`. Highlight `Plot 1` and make sure there are no checks next to any other plots. If there are, arrow to the plot and press `F4`.
4. Then arrow back to `Plot 1`.
5. Press `F1: Define`. For `Plot Type`, press the right arrow and press `1:Scatter`.
6. Arrow down, press the right arrow and press `1:Box`.
7. Arrow down to `x`. Press `alpha L1`. Arrow down to `y` and press `alpha L2`.
8. Arrow down to `Use Freq and Categories?` and use the right arrow. Highlight `NO` and press `Enter`.

9. Press `Enter` again.
10. Press `F5 Zoomdata`. You should see the scatterplot. Press `F3` to trace and the arrow keys to see the coordinates of the points.

Calculating the Regression Equation

1. Press `APPS`, `1:Flashapps`, and `Enter`.
2. Press `F4:Calc`.
3. Press `3:Regressions`.
4. Press `1:LinReg (ax+b)`. For `x List`, enter `alpha L1`. Arrow down. For `y List`, enter `alpha L2`.
5. For `Store RegEqn to:`, arrow right and arrow down to `y1(x)` (or any one of the y's) and press `Enter`.
6. Press `Enter`. You should see a screen with `a`, `b`, `r^2`, and `r` on it. The regression is complete.
7. Write down the equation from the information and press `Enter`.
8. The linear regression is `yhat = 11.2267 - .2133x`

Drawing the Regression Line

1. Press the key with the green diamond on it (it is below the `2nd` key) and press `Y=` (above the `F1` key).
2. This is the **Line of Best Fit**.
3. Press the key with the green diamond on it and press `GRAPH` (above the `F3` key). The line will be drawn.

Outliers

1. When you do Linear Regression, a list called "`resid`" is created automatically.
2. This list has the "`y – yhat`" values in it.
3. Go into `Flashapps` and into your lists.

4. Arrow to "`resid`" so you see it.
5. Then, arrow to the list name `L3`. Press `F3`.
6. Press `1:names`. Arrow down to `STATVARS`. If it is not checked, press `F4`.
7. Then, arrow down to `resid`. Press `Enter`. Press `^2`. Press `Enter`. `L3` will have the `(y - yhat)^2` values.
8. Press `HOME` and Press `F1 8` to clear it.
9. Press `2nd MATH`. Press `3:List`. Press `6:Sum`. Press `alpha (L3)`.
10. Depending on how you have `MODE Display Digits` set, you should see approximately 137.1467. This is the SSE.
11. Calculate s. (You should be `HOME`.) Press `clear`.
12. Press the square root symbol and enter `137.1467/2`. (You get the denominator by taking the number of data points and subtracting 2: $4 - 2 = 2$.)
13. Press `Enter`. You should see 8.2809 (to 4 decimal places).
14. Press the times key and enter `1.9`. Press `Enter`. You should see `15.7337`.
15. Press `clear`. Press `2nd VAR-LINK`.
16. Arrow down to `resid` (it is below `STAT VARS`). Press `Enter`. Press `Enter` again.
17. Arrow up to the list. Scroll through the list using the arrow keys.
18. Compare 15.7337 to the absolute values of the numbers in the list. If any absolute value is greater than or equal to 15.7337, then the corresponding point is an outlier.
19. Absolute values of the numbers in the list are approximately 5.59, 1.95, 2.31, 9.84.
20. None of them are greater than or equal to 15.7337.
21. Therefore, for this data set, no point is an outlier.